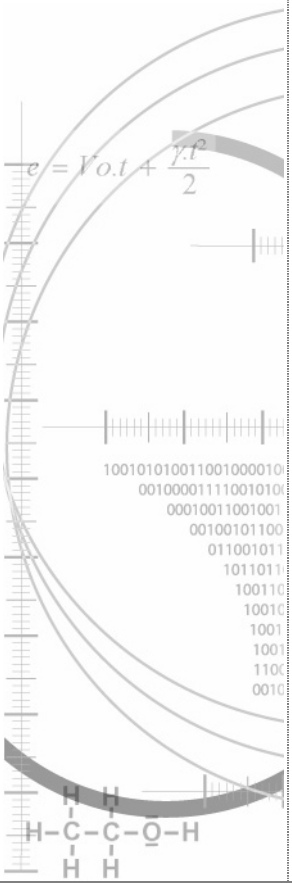
Virtual Machines and Entries Management for a Virtual Machine Logbook in ATLAS

ACRONYM	VM Atlas 2
APPLICANT	Lawrence Berkeley National Laboratory, USA
STUDENT	Julien Poffet
PROFESSORS	Ottar Johnsen, Frédéric Bapst
EXPERT	Peter Kropf
N°	D08I13
TYPE	Diploma project
CONTACT	P. Calafiura



ATLAS is both one of the detector in the LHC particle accelerator at CERN, and the related software (organized in a thousand packages) used by physicists when writing data analysis programs. Sharing the work among the large ATLAS community can be hard, because the different working environments can lead to different behaviors. In this context, virtualization appears as a promising technology, enabling the exchange of the full context of a particular program run. But this raises new questions around the management of numerous virtual machines among the ATLAS community.

This leads to the idea of the Virtual Machine Logbook (VML), an application whose goal is to simplify the share of the work environments used in the ATLAS experiment. The logbook will be a repository which stores different environments used by physicists. VML entries represent virtual machines that can be added, checked out, or removed, thus helping managing the history of the work, and the sharing between users. Two EIA-FR students will develop VML at Berkeley, during a 30-weeks period.

This particular project deals with VML entries management, focusing on the following tasks:

- interfacing the virtualization platforms
- repository management
- command line interface
- VML entries import, export, and transfer
- VM deployment & management (if time permits)

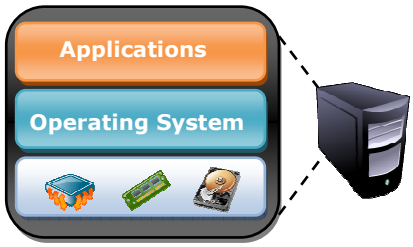
Introduction

Virtual Machine and Entries Management is a part of the Virtual Machine Logbook (VML) project. The second part of this project is developed by another EIA-FR student: Andrea Cavalli. We are both working at the Lawrence Berkeley National Laboratory (LBNL) in California. The project is still in development and will end in February 2009. Also this abstract presents the VML development state and the further work.

Context

There is already a project at CERN called CernVM which helps the physicists in their work by using the benefits of the virtualization. The idea of CernVM is to make the physicist working inside a virtual machine (VM). A VM is a software simulation of a real computer which can hold its own operating system and applications. A VM run as an application on a real computer.

The advantage of using a VM is that both the user data and the context remain inside the VM. Then the physicist shares their VMs instead of sharing their programs. By doing this, they exchange the full program's context.

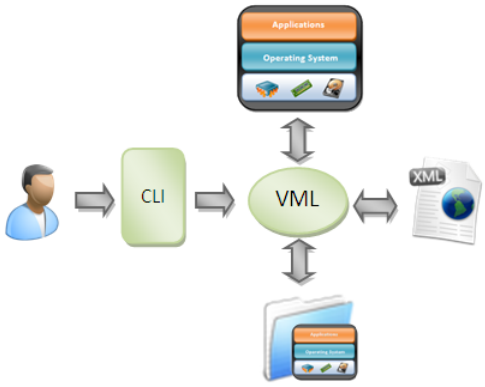


Virtual Machine Logbook

VML is a tool to ease the day-to-day work of the physicist with their VMs. It allows them to easily organize and log their work on multiple projects, speeding up "context switches" from one project to another, and to share their work with other physicists. Since the different virtualization platforms have no inter-compatibility, VML will provide a technology-independent interface to the physicist. VML will store the physicists' work in the most efficient way as possible in terms of disk and CPU usage to make practical to use it ~~also~~ on current generation laptops.

Architecture

The physicists interact with VML by typing commands in a shell. This mechanism is called command line interface (CLI). The CLI of VML is used to parse the command typed by the user and perform the related actions. Then VML use a component called repository. The repository is the place where VML saves the backups of the physicist's work. It keeps a history of the work by maintaining up-to-date an XML file. The repository also constructs a couple of data



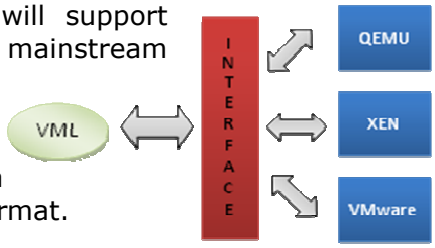
structure based on its XML file to VML. These structures ease the way how VML search and store the backups and VMs meta-data.

Work sharing

One of the goals of VML is allow the physicist to share their work in a simple and efficient way. To meet theses requirement, VML offers the export and import capabilities. Since every physicist work with a CernVM virtual machine, VML will not transfer the whole VM but only the changes that the physicist has perform inside it. Thanks to this, VML optimize the size of the data to transfer.

Technology independent

At this point, VML support VMware virtual machines. The further work is to add to VML an interface which will support the other mainstream virtualization platforms. VML will also keep the VM in a cross-platform format.



Deployment

If time permit VML will allow the physicists to directly submit their work on a cluster/cloud. A cloud/cluster is a center composed by many computers in order to give a lot of computational power.

Conclusion

Working at the LBNL on a project which is part of a world-wide experience like the LHC particle accelerator at CERN is of course rewarding and hugely interesting. VML is a promising project and it will be great if it would be distributed as a tool with the CernVM virtual machine.